MMLA Approach to Track Collaborative Behavior in Face-to-Face Blended Settings

Pankaj Chejara, Reet Kasepalu, Shashi Kant Shankar, Luis P. Prieto, María Jesús Rodríguez-Triana
Tallinn University
pankajch@tlu.ee, reetkase@tlu.ee, shashik@tlu.ee, lprisan@tlu.ee, mjrt@tlu.ee

Adolfo Ruiz-Calleja
University of Valladolid
adolfo@gsic.uva.es

ABSTRACT: Collaborative learning is a complex and multifaceted phenomenon which requires teachers to pay close attention to their students in order to understand the underlying learning process and to offer needed help. However, in authentic settings with multiple groups, it becomes extremely difficult for teachers to observe each group. This paper presents our current MMLA prototype, which allows the collection, analysis and visualization of two types of data from students: audio and logs. We showcase our idea using a Raspberry Pi-based prototype (named CoTrack) for capturing and understanding the students' behavior during face-to-face blended collaborative learning situations. More specifically, CoTrack captures audio data together with software logs captured from their activities using a digital tool Etherpad. Later on, the collected data collected is analyzed to extract the participation behavior across physical and digital spaces. CoTrack has been used in 2 lab and 2 authentic case studies. Preliminary results show that despite of manual set-up and accuracy problems which may emerge, practitioners have shown interest in using it in their (authentic) classroom practice.

Keywords: Multimodal Learning Analytics, Collocated Collaboration

1 INTRODUCTION

Multimodal Learning Analytics (MMLA) has offered a new perspective for understanding learning by utilizing a wide range of sensors and machine learning algorithms (Ochoa, 2017). In addition to informing how learning takes place in real-world settings, MMLA can also “generate distinctive insights into what happens when students create unique solution path to problems, interact with peers, and act in both the physical and digital space” (Blikstein & Worsley, 2016). Researchers have demonstrated the usefulness of MMLA in understanding a range of learning constructs, e.g., emotion, attention, level of expertise, collaboration behavior, and cognition (Di Mitri, Schneider, Specht, & Drachsler, 2018). However, the deployment of MMLA in authentic learning settings is extremely difficult due to the challenges of multimodal data collection and analysis (such as the complex technological set-up, multimodal data fusion, or noisy data) (Chua, Dauwels, & Tan, 2019). These issues need to be addressed in order to raise MMLA adoption in authentic settings.

Collaborative learning in face-to-face (F2F) blended settings includes usage of digital collaboration tools with F2F interactions. However, researchers have either focused on F2F or digital interactions to understand collaboration behavior, but not much work has addressed these two spaces together (Rodríguez-Triana et al., 2017). Thus, showing the participation behavior in digital and physical spaces could potentially be helpful for practitioners and researchers to understand collaboration among students. However, the individual analysis of the spaces poses certain limitations: since digital contributions from the students are not taken into account in F2F interaction analyses and, viceversa, log-based LA tools miss F2F interactions. Such lack of joint analyses is in part justified by the multimodal data collection and analysis challenges (e.g. synchronization and fusion).

This paper presents an MMLA prototype -CoTrack- for data collection, data analysis, and visualization to understand collaborative learning in F2F blended settings. Concretely, CoTrack captures students' interactions from F2F discussions and written tasks through audio data and logs, respectively. Once collected, interactions from both spaces are mapped to the corresponding students in order to measure their participation (e.g., speaking time and number of edits). Finally, different visualizations are generated to enable post-hoc reflection by practitioners.

2 DATA COLLECTION

Researchers consider talk the most important resource in collaboration (Roschelle & Teasley, 1995). In fact, audio features e.g. verbal (speech) and non-verbal (pitch, energy) features are good predictors of collaboration quality and success (Praharaj, Scheffel, Drachsler, & Specht, 2018). Also, when collaboration takes place through digital means, user interactions have been extensively used to understand collaboration behavior. These findings led us to capture both audio data and digital traces. The another rationale for restricting the prototype to audio and logs is to make prototype and its deployment simple and cheaper (as its target is eventually wide authentic settings deployment).

Our idea for the research prototype is motivated by the work of (Noel et al., 2018) which explored the collaboration behavior during collaborative writing activities. This work used the Raspberry Pi module with Microphone array to capture audio data during collaborative writing. Their work focused on F2F interactions by capturing audio data during the collaborative writing and generated visualization (e.g. social network). Our prototype, however, considers digital logs as well collected from students' writing activities in Etherpad (collaborative tool). We developed a similar prototype using Raspberry Pi[1] (3 Model B+) (Learning, 2016) and 4-Mic Microphone array (ReSpeaker) to capture social interaction pattern through audio data. In addition, we developed a plugin to collect students' interactions in a real-time collaborative editor tool: Etherpad. Figure 1 offers an overview of our prototype for capturing the multimodal data during collaboration activity.

For preprocessing, CoTrack uses VAD (Voice Activity Detection) and DoA (Direction of Arrival) algorithms (from library shipped with microphone) to associate the captured audio data with the corresponding student (as each student sits at a particular degree around CoTrack). This data contains

[1] https://www.raspberrypi.org/products/raspberry-pi-1-model-b-plus/

direction (in degrees) from which the sound is detected for every 200 ms duration. Each student is represented by an alias name (e.g.,user-1,user-2,user-3,user-4). To map the Etherpad logs to students, we collect IP addresses before the collaboration activity. Later, CoTrack extracts features such as speaking time and sequence of who spoke after whom. These measures are computed for different time windows (e.g. 2 min, 5 min, 15 min). From Etherpad logs, two features are extracted by the current version of the prototype: number of characters added and number of characters deleted. In addition, features from Etherpad logs (e.g. number of chars added or deleted, text), are merged with audio features (e.g. speaking time, speaking turns). These extracted features are then stored in a database for the purpose of analysis.

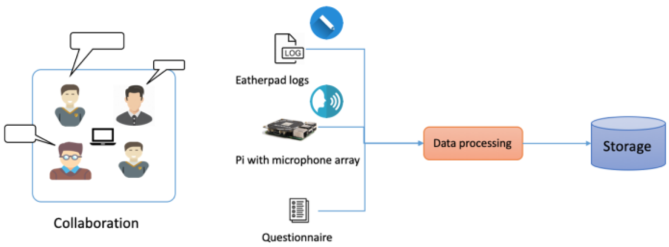


Figure 1: Data collection using CoTrack

3 DATA ANALYSIS

In the analysis phase, we perform an exploratory analysis addresses the following research questions:

i. Which features from collected data are good predictors of collaboration?

ii. How the collected multimodal data can be used to understand the participation behavior?

iii. How useful are the generated multimodal data visualizations for understanding participation behavior?

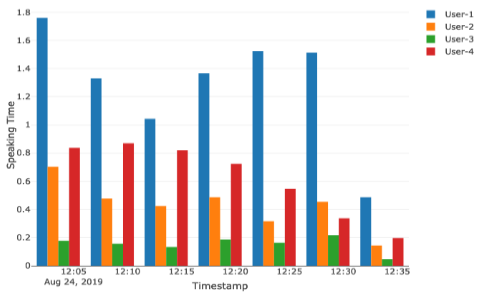
To address these research questions, we developed an algorithm to process and visualize the activity traces and conducted a semi-structured focus-group interview with the teachers.

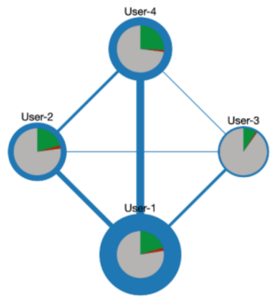
4 DATA VISUALIZATION

In this phase, we visualize the participation behavior captured through audio data and digital traces (Etherpad logs). Particularly, the current version of the prototype shows the overall speaking time, speaking time for different time window, interaction network, and number of characters added or deleted in Etherpad tool for each student. Figure 2.a shows the visualization for speaking time for the time window of 5 mins. This visualization is generated from the datasets collected from one of the experiments reported in Table 1. Each student is shown with a different color.

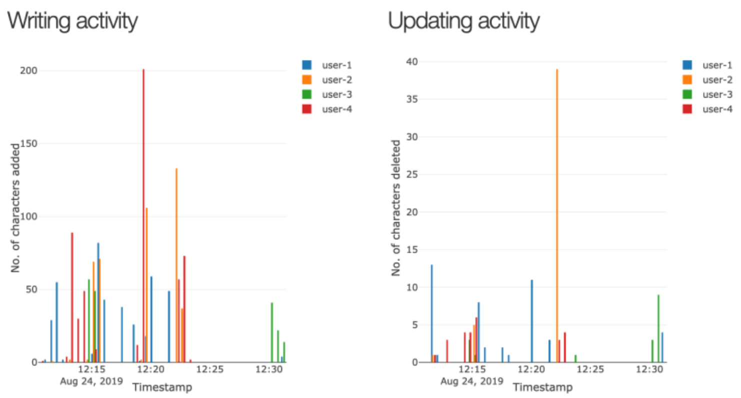
Figure 2.b shows the overall group interaction network with their Etherpad activities. This network is generated from the speaking sequence which is basically a sequence of "who spoke after whom". Each

student is represented by a node and the edge represents the interaction between students. The thickness of the edge shows the frequency of interaction. Additionally, the width of circle outer line represents the speaking time and percentage of characters added or deleted by each student shown by pie chart in the node (e.g. green: % chars added, red: % of chars deleted, grey: % of chars added or deleted by others). We also visualize features (e.g. number of edits) extracted from Etherpad logs for the entire duration of the activity to offer participation behavior in digital space. An example is shown in Figure 2.c.





a. Speaking time per user **b. Group overall interactions with**



c. Etherpad features per user

Figure 2: Visualizations

5 PRELIMINARY RESULTS

CoTrack has been used in two labs and two authentic settings. Table 1 shows the characteristics of those settings. In one of the lab settings, the practitioner herself conducted the data collection process using the web-interface of CoTrack.

Table 1: Characteristics of settings.

Settings	Group-size	Groups
Lab	4	1
Lab	4	1
Classroom	3	3

Classroom	4	2

5.1 Data Collection Protocol

In both lab and classroom settings, we first setup the CoTrack for each group and powered them up. Participants are requested to sit in a particular manner around the prototype (e.g. first participant at 45 degrees, second at 135 degrees, and so on). Then, the server machine is synchronized with the NTP server (running on one of the Pi). Once the technical infrastructure is ready, we provided a brief introduction about the prototype and the purpose of data collection. After getting the written consent for data collection from the participants, we started the Etherpad server and provided the instructions to access it on their laptops. In the classroom settings, due to time constraints, we setup the Etherpad access on each laptop before the activity. We collected the IP addresses of each laptop to map it to the corresponding participants Once everyone had access to Etherpad, we started the audio recordings using CoTrack. For the ground truth purpose, we also video recorded the sessions. Once the activity was finished or the teacher notified about the end of activity, we stopped the audio and video recordings. Finally, we generated visualizations of collected data and showed it to the participants/teacher after the activity.

5.2 Initial Results

The current version of CoTrack utilizes only DoA data to compute speaking behavior, hence, our first aim was to investigate the feasibility of DoA. We manually annotated one group's (from authentic setting with group-size four) audio recording with speaker label, and compared it with CoTrack's results. For this comparison, we only considered annotation frames where only one participant was speaking because the CoTrack can not detect overlapping speaking activity. We determined accuracy by computing the percentage of frames (at the level of 200ms) correctly detected by CoTrack for each participant. The overall accuracy was 48%. We manually checked the video recordings to find out the reason of the low accuracy. We found that frequent moving of participant-3 towards participant-2 during the activity caused the issue of wrongly detecting audio from participant-3 as coming from participant-2. Finally, we found that sitting arrangement and movement of participants can influence the accuracy measure. Additionally, audio noise can also degrade the quality of collected DoA data.

6 CONCLUSION AND FUTURE WORK

In this paper, we presented CoTrack, an MMLA prototype for data collection, analysis and visualization of F2F collaborative learning activities. During the workshop, participants will be able to try the prototype, discuss about its pros, cons, and potential improvements, as well as learn how it could be adapted to their own CrossMMLA contexts. It will also help participants to see its benefit in understanding the social aspect of collaboration with automated data collection and analysis. In future stage of this research, we plan to use questionnaire data and collaboration quality rating schemes for collaboration measure to identify types of collaboration patterns and corresponding multimodal features.

REFERENCES

Blikstein, P., & Worsley, M. (2016). Multimodal Learning Analytics and Education Data Mining: using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3(2)*, 220-238.

Chua, Y. H. V., Dauwels, J., & Tan, S. C. (2019, March). Technologies for automated analysis of co-located, real-life, physical learning spaces: Where are we now?. In Proceedings of the 9th International Conference on Learning Analytics & Knowledge (pp. 11-20). ACM.

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning, 34(4),* 338-349. https://doi.org/10.1111/jcal.12288

Learning, U. (2016). Raspberry Pi 3: *Get Started with Raspberry Pi 3 a Simple Guide TO Understanding and Programming Raspberry Pi 3 (Raspberry Pi 3 User Guide, Python Programming, Mathematica Programming).* CreateSpace Independent Publishing Platform.

Noel, R., Riquelme, F., Mac Lean, R., Merino, E., Cechinel, C., Barcelos, T. S., ... & Munoz, R. (2018). Exploring Collaborative Writing of User Stories With Multimodal Learning Analytics: A Case Study on a Software Engineering Course. IEEE Access, 6, 67783-67798. https://doi.org/10.1109/ACCESS.2018.2876801

Ochoa, X. (2017). *Multimodal Learning Analytics*. Handbook of Learning Analytics, 129-141. https://doi.org/10.18608/hla17.011

Praharaj, S., Scheffel, M., Drachsler, H., & Specht, M. (2018, September). Multimodal analytics for real-time feedback in co-located collaboration. In European Conference on Technology Enhanced Learning (pp. 187-201). Springer, Cham.

Rodriguez Triana, M. J., Prieto Santos, L. P., Vozniuk, A., Shirvani Boroujeni, M., Schwendimann, B. A., Holzer, A. C., & Gillet, D. (2017). Monitoring, awareness and reflection in blended technology enhanced learning: a systematic review. *International Journal of Technology Enhanced Learning*, *9*(ARTICLE), 126-150.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In *Computer supported collaborative learning* (pp. 69-97). Springer, Berlin, Heidelberg.